DOCUMENT RESUME

ED 127 265 SP 010 269

AUTHOR Gall, Meredith D.
TITLE The Importance of Context Variables in Research on Teaching Skills.
PUB DATE 76
NOTE 9p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Curriculum Design; *Educational Research; *Effective Teaching; Elementary School Teachers; Grade 6; Instructional Materials; *Questioning Techniques; Students; Teaching Methods; *Teaching Skills

ABSTRACT

In two experimental studies, the effectiveness of teaching skills are examined in the context of: (1) a specific teaching method linked to (2) specific curriculum objectives and (3) specific curriculum materials; the method and materials used are by (4) specific teachers and (5) specific students in a (6) specific instructional setting; finally, (7) tests are created that are responsive to the idiosyncratic outcomes of the teaching skills and curriculum objectives. Problems and possible solutions involved in creating such a context for studying teaching skills are examined. By referring to these experiments on how the questioning skills of sixth-grade teachers affect student learning outcomes, an example is provided of how methodological requirements are met with respect to each of the seven context variables. Each context variable is considered separately, and the three following methodological requirements are satisfied: (1) ecological validity, the creation of each aspect of the experimental context so that it reflects real classroom teaching; (2) fidelity of treatment, a close correspondence between intended context and what actually occurs experimentally; and (3) achievement of experimental control, either by holding context variables constant or by allowing them to vary randomly. A major problem cited is the expense of conducting the experiment since resources were not already available to design and run teaching skills experiments on this scale. It was decided that in the future research costs can be reduced by designing experiments that use instructional content developed in previous studies. (MM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

The Importance of Context Variables in Research on Teaching Skills[1]

Meredith D. Gall
University of Oregon[2]

ED127265

How effective are teaching skills -- for example, the use of verbal praise, acknowledgment of student ideas, and higher cognitive questions? Do teachers who use these skills promote better learning outcomes in their students than do teachers who use these skills to a lesser degree, or not at all? There has been much interest recently in doing experimental studies to determine the effectiveness of teaching skills (Dunkin and Biddle, 1974).

In doing these studies, researchers need to take into account the fact that teaching skills do not exist in a vacuum. They must be studied in the context of a specific teaching method; the method must be linked to specific curriculum objectives and perhaps to specific curriculum materials. Furthermore, the method and materials must be used by specific teachers and students in a specific instructional setting. Finally, tests must be created which are responsive to the idiosyncratic outcomes of the teaching skills and curriculum objectives.

If a meaningful context is not created, researchers will have difficulty interpreting the results of experiments on teaching skills. The purpose of this paper is to describe problems and possible solutions involved in creating a context in which teaching skills can be studied experimentally. Each context variable will be considered separately: teaching method, curriculum objectives and materials, teachers, students, instructional setting, and testing procedures.

Three methodological requirements must be satisfied in creating each aspect of the experimental context. First is ecological validity, that is, the creation of each aspect of the experimental context so that it reflects "real" classroom teaching. The second requirement is fidelity of treatment, that is, a close correspondence between the intended context and what actually occurred when the experiment was "run." The third methodological requirement is to achieve experimental control either by holding context variables constant or by allowing them to vary randomly.

The methodological requirements with respect to each context variable can be analyzed by referring to two experiments on questioning skills conducted by the author and his colleagues (Gall et al, 1976). The purpose of the two experiments was to determine student learning outcomes which are affected by sixth-grade teachers' use of certain questioning skills. In the first experiment teachers conducted discussions which differed in the presence or absence of two questioning skills: probing and redirection. Probing occurs when the teacher seeks to improve the quality of a student's initial answer to a question by asking a follow-up (that is, a "probing") question. Redirection occurs when the teacher calls upon more than one student to respond to a question already asked of another student. It was hypothesized that probing and redirection would promote learning since they provide students with practice in organizing their facts and ideas into overt responses.

In the second experiment teachers conducted discussions which differed in percentage of higher cognitive questions (25%, 50%, or 75%). A higher cognitive question was defined according to two criteria: first, it must require the student to state predictions, solutions, explanations, evidence, generalizations, interpretations, or opinions. The second criterion is that the prediction, solution, etc. asked for in the question must not be directly available in the curriculum materials. It was

SP010 269

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

hypothesized that discussions with a high percentage of these questions would promote more learning than discussions with a low percentage of these questions.

Teaching Method

The questioning skills in the two experiments were studied in the context of the discussion method. Originally the research team considered creating systematic variation by asking some teachers to conduct discussions in which they made use of certain of the skills, such as probing and redirection. Another group of teachers (or possibly the same group of teachers) would be asked to conduct discussions in which they omitted use of these skills.

This approach creates several problems. First, many features of the discussion which might affect learning outcomes are left uncontrolled. For example, some teachers might ask more questions related to the curriculum objectives than other teachers might ask; or they simply might ask a greater number of questions. As a result, some students might have more "opportunity to learn" than others. If teachers in the experimental group (use of probing and redirection) happened to ask more questions not related to the curriculum than the control teachers (absence of probing and redirection), the potential positive effects of the skills might be cancelled.

Another possible problem with this approach is that teachers within each treatment group might implement their treatment to varying degrees. For example, some teachers in the experimental group might fail to use probing and redirection, whereas some teachers in the control group might use these skills inadvertently. If this phenomenon occurs, it becomes very difficult to determine just what the "treatment" was.

These problems create a need for experimental control of the teaching method in which the skills are embedded. The solution developed by the author and his colleagues was to create a format designated "semi-programmed discussion." The discussions were programmed in that they were scripted in advance by the research team. The teachers were instructed to follow the scripts exactly. Each scripted discussion contained sixteen questions in a fixed sequence. This aspect of the discussion method could be specified rather exactly, thus insuring uniformity of treatment across teachers. However, probing and redirection could not be scripted to the same degree. For example, if a student whom the teacher called upon to answer a higher cognitive question gave a good answer, the teacher would have little justification to probe. In contrast, a weak answer would justify probing. Because of this problem, teachers were trained to use general guidelines rather than exact scripts with respect to probing and redirection. Hence, the designation, "semi-programmed" discussions.

The semi-programmed format has the advantage of experimental control, but it is unrepresentative (low ecological validity) of classroom teaching since teachers typically do not structure discussions to this extent. However, the researchers attempted to build representativeness into other features of the experimental discussions. For example, pilot work by the research team indicated that sixth-grade teachers typically conduct discussions of 20 to 30 minute duration, during which they ask fifteen to twenty substantive questions. The semi-programmed discussions contained sixteen questions in order to represent the natural setting.

The scripts were field-tested to assess the feasibility of the semi-programmed format. Results of the field test indicated that teachers, if trained, could follow the scripts in a natural manner. The probing and redirection treatment generated substantial discussion for most of the questions; the no probing and redirection treatment could be likened to a simple question-and-answer format. Some teachers commented that the scripts facilitated the teaching task: since they knew in advance the questions to be asked, they could devote more attention to the students' behavior. The research team and other persons listened to audiotapes of the semi-programmed discussions. The discussions appeared natural and realistic to them. At the conclusion of the experiment, students generally expressed positive attitudes on a Likert-type scale toward the semi-programmed discussions.

These results demonstrate the feasibility of using a programming technique to achieve experimental control over teaching method as a context variable, while at the same time preserving representativeness.

Another advantage of the semi-programmed format is that it promotes fidelity of treatment. In the two experiments on questioning skills, teachers were asked to audiotape several of their discussions. The sampling procedure was used instead of audiotaping all discussions because teachers found the procedure somewhat disruptive. Raters were employed to score the tapes on variables considered to be critical to treatment fidelity. For example, the raters checked each discussion to determine whether the teachers (a) asked or omitted each scripted question, (b) made substantial changes in the wording of a question, and (c) asked the questions in a different order than was presented in the script. Results of this analysis indicated that deviations from the scripts were infrequent. The most deviance occurred in the order of question-asking. Deviations were observed in only six of the 144 audiotaped discussions.

The raters observed frequency of probing and redirection in each treatment group. As intended, teachers in the probing and redirection treatment followed the guidelines by using these skills with high frequency: probing ($\bar{X}$ = 21.45) and redirection ($\bar{X}$ = 22.96). In the no probing and redirection treatment, teachers were instructed to avoid use of these skills. Audiotape analysis revealed high, though not perfect, fidelity of treatment: probing ($\bar{X}$ = 3.29) and redirection ($\bar{X}$ = 0.33). Other analyses also revealed high fidelity of treatment.

In summary, it appears that the semi-programmed format is very useful in helping teachers, first, to use the teaching method intended by the researchers, and, second, to make a quantitative check of the correspondence between the researchers' intent and the teacher's behavior, that is, fidelity of treatment.

Curriculum Materials and Objectives

Curriculum objectives and materials form another important set of context variables which should be controlled in research on teaching skills. In the author's experiments, questioning skills were studied in the context of the discussion method, which in turn was used in conjunction with a specially prepared set of curriculum materials. Each experimental lesson consisted of a scripted discussion preceded by a brief (approximately 15 minutes) curriculum selection which students viewed and/or read. The same curriculum materials were given to all students in the

experiment, thereby equating opportunity to learn the curriculum across treatment conditions.

The curriculum materials were constructed to meet the requirements of experimental control and ecological validity, that is, representativeness with respect to classroom teaching. The latter goal was accomplished by including several features in the materials. First, the curriculum consisted of ten lessons, one per school day, for a total of approximately ten hours of instruction. This length is typical of curriculum units at the sixth-grade level. By having a curriculum unit of this duration, the experiments avoided a criticism sometimes directed at other research on teaching skills, namely, that the total teaching/learning episode was too short (sometimes as short as 30 minutes) for effects of the teaching skills to be observed.

The second representative feature of the curriculum unit was that it provided instruction in ecology, a topic which was perceived by the participating schools as relevant and consistent with their objectives. Thus, the curriculum was complementary to the regular school curriculum, rather than an irrelevant intrusion. Third, the ecology materials were constructed by following a typical curriculum development sequence. Objectives for the unit were identified; next reading selections, audio-visuals, and scripted discussions were developed to help students achieve the objectives. Finally, the various components were field-tested and revised to get the "bugs" out. This process was designed to insure that the questioning skills were used in the context of an effective curriculum. Otherwise the ineffectiveness of the curriculum might obscure potential positive effects of the questioning skills under investigation.

The fourth representative feature of the curriculum is that it included a variety of instructional materials over the duration of the unit -- an ecology game, a film, filmstrips, and illustrated reading selections. Variety, which itself has been shown to be effective (Rosenshine and Furst, 1971) contributed to establishing a strong curriculum context in which to investigate the questioning skills. The fifth feature is that the reading level of the materials was kept at or below the students' grade-level so that reading skill would not be a major influence on their ability to participate in the discussions associated with each treatment condition. Finally, the materials were presented in an attractive format to stimulate and maintain students' interest.

Fidelity of treatment with respect to curriculum materials was relatively easy to observe. Teachers were familiarized with the materials and trained in their use prior to the data collection phase. During the data collection phase, the research team visited the teachers' classrooms to determine whether students actually received the materials as intended. Fidelity of this aspect of the treatment was high. In a few instances a teacher forgot to bring the lesson materials to class. When this occurred, the lesson was delayed until the research team monitoring the experiment could provide a back-up set of materials.

In summary, it appears feasible to create a curriculum context which is representative of actual classroom teaching and which is implemented as planned.

Teachers

Teaching skills obviously cannot be studied unless teachers are available to

implement them. Thus, teachers become an important part of the context in which teaching skills are researched. Furthermore, teachers require training if the skills are to be used in the manner desired by the researchers. In the experiments on questioning skills, the training task was simplified by the fact that teachers needed only to learn how to use prepared curriculum materials and a scripted discussion format. They did not need to generate their own lesson plans and materials.

Despite this fact, the experiment still required five days of teacher training initially plus frequent refresher and review sessions throughout the four-week duration of the curriculum units.[3] The decision was made not to use the students' regular teachers, because they would find it difficult to attend the training sessions. Another reason for not using the regular teachers was that the experimental design (a latin square) required four different teachers to work at the same time with a particular class; each teacher administered a different treatment to approximately one-fourth of the students in the class.

Because of the requirements imposed by the research design, a separate group of teachers was trained to implement the treatments. Substitute teachers in a nearby school district were recruited for this purpose. The primary criteria for selection as a teacher in the experiments were: (1) a California teaching credential; (2) experience in teaching elementary school children; and (3) willingness to contribute to a research project. The first two criteria helped to insure that the experimental teachers would be representative of regular classroom teachers. After the initial training sessions, which occurred in schools similar to those used in the experiments, the teachers who were best at following the semi-programmed discussion format and at interacting with children were chosen to conduct the experiments. The success of this procedure was shown by the fact that the experimental teachers were well-accepted by the school administrators and the regular teachers. There was not a single complaint or instance when the experimental treatments were compromised because of friction between school personnel and research personnel.

The use of special teachers also contributed to fidelity of treatment. Because the experimental teachers were paid directly by the research team, they could be held accountable for meeting the requirements of the treatment conditions. The same level of accountability would have been much more difficult to require of regular teachers.

The experimental teachers responded well to the need for accountability. None of them resigned from the project until the experiments had been completed. They maintained a high level of fidelity to the discussion scripts, as described earlier in this paper. Furthermore, the data analyses showed that teacher effects accounted for very little of the variance in the dependent measures, even though the experimental design required each teacher to implement four different treatment conditions. Thus, it appears that teachers can be trained to conform to fairly precise standards of teaching performance.

Students

Students obviously are another important aspect of the context in which teaching skills are used. To investigate questioning skills, especially the skill of asking higher cognitive questions, the research team needed to define a group of students

who were representative of the students with whom these skills are used. Sixth-grade students were chosen for this reason. Discussions involving probing, redirection, and higher cognitive questions are conducted with students in the intermediate (fourth through sixth) grades. Another reason for selecting students at this grade level is that reading materials could be used to present much of the facts and ideas on which the semi-programmed discussions were based. Substantially more effort by the research team would have been required to create curriculum materials which did not involve reading.

The most critical fidelity of treatment issue with respect to students is whether they received their assigned treatment condition. One aspect of this issue is whether students actually attended each treatment session. In the experiments on questioning, attendance was recorded daily and was found to be quite high. Absences occurred only when a student was not in school, usually because of illness. School personnel cooperated with the research by not scheduling individual children for other activities, such as band practice, during the treatment sessions. In a few instances an activity affecting the entire class was scheduled at the same time as the treatment sessions. When this occurred, the treatment sessions simply were rescheduled.

Even when students attend their assigned treatment sessions, it is possible that they may not "receive" the treatment. Students' physical presence does not mean necessarily that they are attending to and processing the curriculum content. Lack of attention is likely to occur if students misbehave or are distracted by another activity. A shortcoming of the experiments on questioning skills is that this aspect of fidelity of treatment -- students' attentiveness -- was not monitored closely. However, informal observation by the research team indicated that the students generally were attentive to the curriculum materials and to the discussion process.

Instructional Setting

In designing the experiments on questioning skills, the research team considered the alternative of bringing students to their institute for the treatment sessions. This procedure would have made it easier to maintain certain experimental controls. However, the institute had the disadvantage of being unrepresentative of the students' typical learning environment. An unfamiliar setting might mask or otherwise alter the effects of the questioning skills being investigated. Therefore, the experiments were conducted at the students' own schools, and whenever possible, in the students' own classroom.

The size of the experimental discussion groups posed another problem. Teachers typically conduct whole class discussions, involving perhaps twenty to thirty students. However, research on group size (Gall and Gall, 1976) suggests that the optimal discussion group is five or six students. The research team decided to create an optimal context in which to evaluate the effects of questioning skills. Therefore, each treatment group consisted of six students. These groups can be viewed as representative of recommended teaching practice rather than as representative of typical teaching practice.

Fidelity of treatment with respect to instructional setting is relatively easy to accomplish and to monitor. In the experiments on questioning skills, the only

problem of note was that the research team occasionally experienced difficulty in locating sufficient room for each of the treatment groups to conduct their designated activity.

Testing

In a curriculum unit lasting two weeks, teachers might administer one or two tests at the end of the unit to measure learning. It is difficult for researchers to conform to this aspect of classroom practice, since they usually wish to administer a variety of tests to measure the many possible effects of teaching skills. In the experiments on questioning skills ten different measures were administered. Six of these measures were administered twice, that is, immediately before and after the curriculum unit. One of the measures was administered three times.

Clearly this amount of testing is low in ecological validity. Yet it appears necessary in order to obtain useful information about possible student effects. To minimize the negative impact of multiple tests, a number of steps were taken: testing was extended over several days; testing periods were kept short, as were the tests themselves; and the tests were printed in an attractive format. Apparently these procedures were effective since the research team rarely received complaints from students or teachers about the testing procedure.

The fidelity issue with respect to testing is whether the tests were administered according to the standardized testing procedures developed by the researchers. The experimental teachers were specially trained in the instructions for administering each test. Also, their performance during the actual test administration was monitored. Students who were absent at the time of group testing were administered the tests individually to preserve completeness of data. These procedures resulted in high fidelity of testing procedures, that is, there was close correspondence between the researchers' intent and what actually occurred during data collection.

Experimental control in this type of research requires that the tests be high in content validity. By content validity is meant the extent to which each test reflects the content (information) taught in the curriculum materials and in the teaching episodes. If a test is low in content validity, it will not be sensitive to the effects of the teaching skill under investigation (Gall, 1974). In the experiments on questioning skills, it was possible to achieve high content validity because the content in the curriculum materials and discussions was prespecified, and therefore could be made available to the test developers.

Concluding Remarks

The preceding discussion stressed three methodological requirements that should be met in designing each aspect of the experimental context in research on teaching: ecological validity, fidelity of treatment, and experimental control. The author's experiments on questioning skills provide examples of how those requirements can be met with respect to each type of context variable.

A major problem with the experiments is that they were very expensive to conduct. Resources usually are not available to design and run teaching skill experiments on

this scale. However, the major costs in these experiments were in the design and development of an instructional context: scripted discussions, curriculum materials, and tests. These costs can be eliminated or reduced in future research by designing experiments which use instructional contexts developed in previous studies.

Footnotes

[1]Paper presented at the annual meeting of the American Education Research Association, San Francisco, April 1976, as part of the symposium, "What does recent research tell us about the study of teaching skills?" I wish to extend appreciation to Dr. Beatrice Ward, who organized the symposium and who suggested the topic of this paper.

[2]The work on which this paper is based was completed while the author was a staff member of the Program for Effective Education at the Far West Laboratory for Educational Research and Development.

[3]The two experiments each incorporated the two-week curriculum unit; hence, the four-week total.

References

Dunkin, Michael J. and Biddle, Bruce J. The Study of Teaching. New York: Holt, Rinehart, and Winston, 1974.

Gall, Meredith D. The problem of "student achievement" in research on teacher effects. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 1973.

Gall, Meredith D. and Gal[illegible] Joyce P. The discussion method. In N.L. Gage (ed.) The Psychology of T[illegible]ing Methods: The Seventy-fifth Yearbook of the National Society for [illegible]e Study of Education. Chicago: University of Chicago Press, 1976.

Gall, Meredith D., Ward, Beatrice A., Berliner, David C., Cahen, Leonard S., Crown, Kenneth A., Elashoff, Janet D., Stanton, George S., and Winne, Philip H. The Effects of Teachers' Use of Questioning Techniques on Student Achievement and Attitudes. San Francisco: Far West Laboratory for Educational Research and Development, 1855 Folsom St., Report A76-2.

Rosenshine, Barak, and Furst, Norma F. Research on teacher performance criteria. In B.O. Smith (ed.) Research in Teacher Education: A Symposium. Englewood Cliffs, N.J.: Prentice Hall, 1971.